A Notion of Prominence for Strategic Games

Alessandro Sontuoso and Sudeep Bhatia

May 24, 2017

Abstract. Identifying the best course of action in games with multiple equilibria is a longstanding unresolved issue in strategic interaction. The concept of prominence as a criterion for equilibrium selection has been suggested, but has remained for the most part an informal notion, without a psychologically grounded characterization. In this paper we propose one such characterization: by drawing on existing theories of human memory, language, and decision making we define prominence in terms of frequency of exposure. In particular, we consider games where strategies are denoted by natural language labels, and we measure the prominence of each strategy by how often its label occurs in natural language corpora. Our specification of prominence yields sharp quantitative predictions about behavior in coordination and discoordination problems. Here we present three studies designed to test such predictions, and show that individuals do select strategies that fulfill our definition of prominence and they furthermore do so in a (boundedly) rational manner.

KEYWORDS: focal points, salience, accessibility, coordination, hide-and-seek, level-k

JEL Classification Numbers: C72, C91.

Contact: Alessandro Sontuoso, Philosophy, Politics and Economics, University of Pennsylvania, 249 S. 36th St., Philadelphia, Pennsylvania, 19104, sontuoso@sas.upenn.edu; Sudeep Bhatia, Department of Psychology, University of Pennsylvania, 425 S. University Ave., Philadelphia, Pennsylvania, 19104, bhatiasu@sas.upenn.edu.
Introduction

Imagine that you have to meet someone in New York tomorrow, but neither of you can communicate about where or when to meet: what would you do? Thomas Shelling (1960) used this question to illustrate the ubiquity of simple coordination problems, the difficulty of equilibrium selection in such problems,\(^1\) and the role of salience or prominence in facilitating coordination. After conducting an informal survey, Schelling found that a response frequently given by his students was “waiting under the clock in Grand Central Station at 12 noon,” evidently a prominent place for his cohort.

Since Shelling’s initial work, prominence has received some attention in the experimental study of coordination (Mehta et al., 1994; Bacharach and Bernasconi, 1997; Crawford et al., 2008). However, the psychological underpinnings of prominence in strategic settings have not been fully explored, and it is unclear how decision processes determine prominence in coordination or discoordination games:\(^2\) for this reason, it is often difficult to make clear-cut predictions about the strategies chosen by individuals in many of those games.

In this paper we aim to address such issues. To that end, we define prominence in terms of \(“freqency\ of\ exposure”\), and so we measure the prominence of strategies by how often their labels occur in natural language. Note that the frequency of exposure to words (such as natural language strategy-labels) is a key determinant of the mental accessibility or fluency of the concepts corresponding to those words, and thus it is closely correlated with observed behavior in memory, language, and decision tasks (Alter and Oppenheimer, 2009; Anderson and Schooler, 1991; Dougherty et al., 1999; Gigerenzer and Goldstein, 1996; Griffiths et al., 2007; Tversky and Kahneman, 1974). For example, locations that are more likely to be mentioned in everyday language are also more likely to be remembered and recognized, as well as more likely to be used in unrelated language tasks; besides, they are more likely to be judged as being large, important, or desirable in simple decisions (e.g., Goldstein and Gigerenzer, 2002). Likewise, lists

---

\(^1\) The class of pure coordination games includes any problem where each player aims to make the same choice as everyone else: each player receives an identical payoff in case of a match; each player receives a payoff of 0 otherwise. For example, in Shelling’s game both players’ strategies are defined by the set of meeting points in New York (we disregard the time for expositional simplicity). Hence, when players select the same meeting point, they each receive a positive payoff; by contrast, when players select different meeting points, they each receive a payoff of 0.

\(^2\) A discoordination problem is represented by any 2-player game with a common strategy set, where one player aims to match the counterpart’s strategy while the other aims to avoid any such match. A classic instance of a discoordination problem is given by the Hide-and-Seek game: we specify its payoff structure below.
consisting of famous, frequently mentioned names are judged to be larger than lists consisting of relatively obscure names (e.g., Tversky and Kahneman, 1973), and causes of death that are frequently mentioned in news media are rated to be more probable than others (e.g., Lichtenstein et al., 1978).

In the context of strategic interaction – and coordination games in particular – those findings suggest a straightforward hypothesis: subjects may perceive as particularly prominent those strategies that are denoted by labels (i.e., words) frequently used in natural language; we further hypothesize that such strategies are more likely to be selected in coordination games. Our hypotheses are motivated by the following psychological mechanism. First, due to the close link between exposure and mental accessibility, one can easily identify which strategy-labels are more prominent in a set of \( m \) words, given one’s everyday experience with those \( m \) words.\(^3\) Second, if players “project” their mental categories onto others,\(^4\) they will realize that others too can identify the same, prominent strategy-labels (on condition that players come from the same cultural and linguistic environment). Third, since the first two steps imply almost no cognitive burden, then relying on frequency of exposure to guide one’s decisions turns out to be an efficient decision rule (in terms of time and effort) for coordination problems.\(^5\)

Note that – from the viewpoint of the researcher – our proposed notion of prominence as frequency of exposure has the advantage of yielding clear-cut, testable predictions. In fact, the frequency of exposure in natural language can be easily quantified with the use of corpus analysis,\(^6\) which in turn can be used to predict responses in games with natural language labels. This is a relatively unique feature of our frequency-of-exposure definition of prominence, as most theories of salience do not lend themselves to sharp quantitative predictions. In what follows we present three studies devised to test the implications of our definition.

\(^3\) See Herzog and Hertwig (2013) for an overview of the desirable properties of frequency of exposure, and related memory-based cues, in (non-strategic) decision making tasks.

\(^4\) For an early account of projection effects, see Ross et al. (1977).

\(^5\) We stress that this kind of shared cultural knowledge facilitated by frequency of exposure is an implicit feature of Shelling’s (1960) definition of prominence. Schelling noted that people often solve coordination problems by resorting to contextual cues – which are likely to be recognized by individuals with the same experience – thereby making one of the feasible strategies prominent. Sugden followed in Schelling’s wake by noting that such cues may depend in part on psychological and cultural factors; in fact, to our knowledge, Sugden was the first to explicitly suggest that players may distinguish between labeled strategies in terms of the frequency with which they have heard them (Sugden, 1995: pp. 547-548).

\(^6\) See Dougherty et al. (2008), Goldstein and Gigerenzer (2002), and Schooler and Hertwig (2005) for relevant examples in the study of judgment and decision making.
We designed our first study to predict behavior in simple “matching games”. For the purposes of this study we considered games featuring two different types of strategy-labels in turn. More specifically, our games featured three options, which were denoted by either names of countries or names of ingredients. In both cases we found a strong correlation between the choices of participants and the frequency of exposure to the strategy-labels, where the latter is measured by the labels’ frequency of occurrence in the Google Books corpus (i.e., the “NGRAM” metric, defined below).

Our second study was designed to verify whether subjects use “prominent labels” (in the sense of the above definition) strategically rather than naively, such as an automatic response. Note that in the latter case a subject would be blindly inclined to select prominent labels, whether she has the incentive to match her counterpart’s choice or not. Hence, we contrasted the behavior of a subject participating in a (i) matching game with the behavior of subjects facing three alternative scenarios, with each scenario featuring exactly the same list of labels. Specifically, we considered: (ii) the case in which a subject is prompted to pick an option, without the explicit objective to match her counterpart’s choice; (iii) the case in which a subject is prompted to avoid matching her counterpart’s choice, under the assumption that her counterpart instead wants to match; (iv) the case in which a subject is prompted to match her counterpart’s choice, under the assumption that her counterpart instead wants to avoid any such match. It should be stressed that – since all four conditions involve exactly the same option triplets – if the effect of the frequency of exposure from Study 1 were merely due to an automatic (naïve) response, then we should observe similar choice distributions across the four conditions. In brief, our data show that participants in problem (ii) (“Pickers”) were about as likely to select the most prominent label as were participants in the matching game; by contrast, participants in problem (iii) (“Hiders”) were less likely to select the most prominent label than were participants in problem (iv) (“Seekers”). As will soon be clear, this pattern suggests a boundedly rational use of strategy-labels.

Finally, we designed our third study to corroborate our results on the strategic use of strategy-labels in matching games: to that end, we varied the cultural/linguistic makeup of the

---

7 By “matching games” we refer to the subclass of pure coordination problems containing any 2-player game, whereby each player faces the same (finite) list of strategies and each is perfectly indifferent as to which equilibrium occurs. In fact, note that in a matching game with $m$ strategies there are $m$ strict Nash equilibria; also note that there players and strategies – as well as equilibria – are perfectly symmetric, which is the reason why the equilibrium selection problem is particularly pressing in such games.
subject pool as well as participants’ knowledge of the counterpart’s country of residence (and thus cultural/linguistic background). Our data show that subjects were less likely to rely on the frequency of exposure as a means to guiding their behavior whenever they knew that their counterpart resided in a different country. Put it differently, when subjects knew that their assigned partner was alike (in terms of cultural background), they were more likely to select the label most prominent in their own vocabulary.

To conclude, our studies provide very robust evidence indicating that individuals do select strategies fulfilling our definition of prominence, and they furthermore do so in a (boundedly) rational manner.

**General Procedures**

Our three studies were conducted online between September and November 2016. Participants were recruited through the Oxford-based crowdsourcing platform Prolific Academic. Participation in our study was limited to individuals with a Prolific-Academic approval rate greater than 95%. At the beginning of each study, subjects were informed that every participant would be assigned a partner at random, and that they would not know the identity of their partner or be able to communicate with him/her. No subject was allowed to participate in more than one study. Participants’ responses were incentivized, as outlined in the sections below.

All three of our studies involved coordination (or discoordination) games with natural language labels. The prominence of the labels was quantified by their frequency of occurrence in natural language. More specifically, the prominence of a label was captured by the case-insensitive average yearly NGRAM frequency of the word corresponding to the label in the Google Books corpus, for books published after 2000. These NGRAM metrics were obtained through the Google NGRAM tool (books.google.com/ngrams) in August 2016, shortly before running the studies. In Studies 1 and 2 we used the general English corpus, whereas in Study 3 we used the American English and British English corpora. Google Books NGRAM frequency is a common metric for word popularity in natural language, and has been previously used for linguistic, psychological, sociological, and historical research (Michel et al., 2011).
Study 1

Demographics. The subject pool for Study 1 consisted of 91 US resident individuals. The average participant was 33 years old, and 57% of the subjects were male. All subjects received a 0.5 GBP participation fee, in addition to any additional payoffs resulting from their choices (as described below).

Design. Study 1 involved a set of matching games, with each game featuring a 3-element strategy set. Specifically, each member of a pair received GBP 0.10 if both players chose the same option; each member of a pair received nothing otherwise. For expositional purposes, below we shall denote the set of options by \{X, Y, Z\}. Figure 1 represents the above-described game structure in bimatrix form, whereby the bottom-left and top-right numbers in each cell indicate the payoffs to Player 1 and Player 2, respectively.

```
      X     Y     Z
    X  0.10     0     0
    0.10   0     0

      X     Y     Z
    Y  0     0.10   0
    0.10  0     0

      X     Y     Z
    Z  0     0     0.10
    0     0   0.10
```

Figure 1 – The matching game. The bottom-left and top-right numbers in each cell represent the monetary payoffs to Player 1 and Player 2, respectively.

Subjects played 10 consecutive instances of the above game, with each instance differing from the others only in the names of the three options (as reported in Figure 2 below). Each subject was assigned the same partner for all the (10) games, but no feedback was provided after each game. The order of the games was randomized across subjects. By contrast, the order of the three options (in a given game) was determined randomly prior to the experiment, but was identical across subjects. The order of the options is reported in Figure 2 below (i.e., option X, option Y, option Z).
Note that we ran two versions of the study, whereby Version A’s options consisted of names of *countries* while Version B’s options consisted of names of *ingredients*. Participants were randomly assigned to the two versions (48 subjects were assigned to Version A and 43 subjects to Version B). Average earnings were GBP 0.63 and 0.52 for Versions A and B, respectively (not counting subjects’ participation fees). Given the odd number of subjects in Version B, one participant was assigned two partners (but received compensation for playing with either one, at random); the two partners were treated like any other participant.

![Figure 2](image-url) – The option sets for Study 1. The left and right panels refer to Versions A and B, respectively. The blue, maroon, and green cells below each strategy-label indicate the option with the relatively highest, middling and lowest NGRAM metric, respectively.
**Results.** The average participant in Study 1 chose the strategy associated with the highest, middling and lowest NGRAM metric respectively 45.27%, 39.34% and 15.39% of the time.\(^8\) This distribution differs strongly-significantly from play in the fully mixed equilibrium assigning equal probability to each option (\(N = 91\) obs., \(T^2 = 127.67, p = 0.000\) under a Hotelling’s \(T^2\)-squared generalized means test conducted on the sample of per-subject mean choices). Also note that the fully mixed equilibrium would imply an expected coordination rate of 0.333, whereas the expected frequency of coordination resulting from our sample of (per-subject) mean observations is 0.376.\(^9\) (That rate measures the average *probability that two individuals* – selected at random from the pool of participants in Study 1 – *follow the same rule*, that is, select the option with the highest, middling or lowest NGRAM metric.\(^{10,11}\)

For the purposes of more rigorously testing our hypothesis, we ran alternative-specific conditional logistic regressions (“asclogit”; see McFadden, 1973) that account for the relative frequency of exposure of each option, as well as for the order in which the three options were presented. The asclogit coefficients for the relative frequency of exposure are positive and strongly significant for both Version A (coef. = .5323, \(z = 8.34, p = 0.000\), under a two-tailed asclogit regression conducted on the full sample of individual observations, with standard errors adjusted for clustering on 48 subjects) and Version B (coef. = .1727, \(z = 2.67, p = 0.008\), under a two-tailed asclogit regression conducted on the full sample of individual observations, with standard errors adjusted for clustering on 43 subjects). In a nutshell, this implies that the higher the relative exposure – as measured by our NGRAM metric – the more likely it is for an option

---

\(^8\) More specifically, the average participant chose the option with the *highest* NGRAM metric 45.83% of the time in Version A and 44.65% of the time in Version B.

\(^9\) The sample of (per-subject) mean observations is obtained as follows. First, for each choice of subject \(i\), assign a value of 1 or 0 to indicate if the option with the *highest* NGRAM was chosen or not; hence, take the average value across all the choices of \(i\). Similarly, assign a value of 1 or 0 to indicate if the option with the *middling* NGRAM was chosen or not, and take the average value across all the choices of \(i\). Lastly, assign a value of 1 or 0 to indicate if the option with the *lowest* NGRAM was chosen or not, and take the average value across all the choices of \(i\).

\(^10\) In keeping with previous studies of coordination we focus on the expected coordination rate, as opposed to the actual frequency of coordination (see, among the others, Crawford et al., 2008). In fact, actual coordination rates depend on individual choices and on a purely random element, that is, the random assignment of partners: it follows that – given a relatively small sample size – such a random element is likely to bias coordination rates. «Thus, the actual frequency of coordination has no special significance; it is more appropriate to consider the expected frequency of coordination» (Mehta et al., 1994, p. 663, italics in original).

\(^11\) Formally, the expected coordination rate is given by \(\sum_a f_a \frac{fa-N-1}{N-1}\), where \(N\) is the number of subjects in the study and \(fa \in \{f_H, f_M, f_L\}\), with \(f_H, f_M, f_L\) denoting the frequency of play of the option with the highest, middling and lowest NGRAM metric, respectively.
to be selected, regardless of its position (i.e., regardless of whether the option is presented at the top, center, or bottom of the list). Please refer to the Appendix for a bar graph of the frequency distributions of individual-level choices for each of the 20 games, as well as for an additional regression and extended commentary.

Before proceeding we note that, although Study 1 shows a significant impact of the frequency of exposure to labels, it does not establish whether this effect is strategic or merely the byproduct of an automatic behavioral response. For this reason, in Study 2 below we provide a first test of the ability of subjects to use such frequency of exposure in a strategic manner.

**Study 2**

**Demographics.** The subject pool for Study 2 consisted of 160 US resident individuals. The average participant was 30 years old and 58% of the subjects were male. As with Study 1, all subjects received a 0.5 GBP participation fee, besides any additional payoffs resulting from their choices (as described below).\(^{12}\)

**Design.** Study 2 involved the following four conditions, whereby each condition featured the same (10) option triplets as in Version A of Study 1 (see the left panel of Figure 2 above).

a. **“Coordinate”**: This is a replication of Version A of Study 1, which – among the other purposes – serves to verify the robustness of our earlier results. Participants in this condition were paired with other participants in this condition.

b. **“Pick”**: This condition featured the same (10) option triplets as in the Coordinate condition, except that here participants were asked to merely pick any option they wanted. That is, participants were not assigned a partner nor did they receive any earnings on the basis of their choices (i.e., they only received the standard 0.5 GBP participation fee) and, hence, they had no strategic incentives to select one option over another.

c. **“Seek”**: This condition featured the same (10) option triplets as in the Coordinate condition, except that here the incentive structure reflected the role of Seeker in the Hide-and-Seek game (represented in Figure 3 below). As shown in Figure 3, a Seeker receives

---

\(^{12}\) Average earnings were GBP 0.38 (not counting subjects’ participation fees).
GBP 0.10 if both members of a pair choose the same option; a Seeker receives nothing otherwise. Participants in this condition were paired with participants in the “Hide” condition below.

d. “Hide”: Again, this condition featured the same (10) option triplets as in the Coordinate condition, except that here the incentive structure reflected the role of Hider in the Hide-and-Seek game. As shown in Figure 3, a Hider receives GBP 0.10 if members of a pair choose different options; a Hider receives nothing otherwise. Participants in this condition were paired with participants in the “Seek” condition above.

Note that – since all four conditions involve exactly the same option triplets – if the effect of the frequency of exposure from Study 1 were merely due to an automatic (or naïve) response, then we should observe the same choice distributions across the four conditions. If, however, participants use such frequency of exposure in a strategic manner, then they will be more likely to select the prominent (i.e., the most frequent) strategy-label in the Coordinate and Pick conditions rather than in the Hide and Seek conditions.

![Figure 3](image-url) - The Hide-and-Seek game structure in bimatrix form. The bottom-left and top-right numbers in each cell represent the payoffs to the Hider and Seeker, respectively.

The above hypothesis is motivated by the following psychological mechanism. First, due to the close link between exposure and mental accessibility, one can easily identify which strategy-label is most prominent in a triplet (given frequency of exposure in natural language); this is likely to drive choices in the Pick condition. Second, if players “project” their mental categories
onto others, they will realize that others too can easily identify the same, prominent strategy-label. Third, since the first two steps imply almost no cognitive burden, then relying on such frequency of exposure to guide one’s decisions turns out to be an efficient decision rule (in terms of time and effort) for choices in the Coordinate condition. It is now clear that our informal mechanism implies no behavioral differences between Coordinate and Pick conditions.

On the other hand, we expect fewer choices of the prominent strategy-labels in the Hide and Seek conditions, with the change in choice distributions varying with subjects’ strategic sophistication. Our intuition can be rationalized by one specification of level-k reasoning; note that level-k theories allow behavior to be heterogeneous and assume that each player “type” best responds to some beliefs, with such beliefs being based on simple nonequilibrium models of others (Crawford et al., 2013). More precisely, level-k types anchor their beliefs in a nonstrategic L0 type and adjust them via iterated best responses, so that L1 best responds to L0, L2 best responds to L1, and so on.13 In what follows we list a set of assumptions that a standard level-k model would necessitate in order to rationalize our predictions across conditions.

(i) All participants in the Pick condition (henceforth, “Pickers”) correspond to a nonstrategic L0 type in the Coordinate condition.14

(ii) L0 choices are not uniformly random and, more specifically, one deviation from randomness is that the probability of selecting the most frequent strategy-label is higher than any other label: let’s denote such a probability by q.

(iii) There are no players at L3 or higher.15

Given that, level-k reasoning implies that – in any coordination game – all types above L0 will select the strategy-label that is the modal choice at L0 with probability one. Here, regardless of the actual distribution of types, this implies that subjects will be weakly more likely to select the most frequent strategy-label in the Coordinate condition than in the Pick condition.


14 A nonstrategic “L0 type in the Coordinate condition” represents an individual who can easily identify which strategy-label is most prominent in a triplet (given her everyday exposure to those labels), but does not have a theory of mind.

15 Some level-k models assume that there are actually no players at L0, and still L1 types naïvely best respond to L0. In this regard, we note that the assumption that there are no players at L0 is not universal; in fact, relaxing such an assumption allows for the possibility of errors (see Crawford et al., 2013).
Also, level-$k$ theories typically assume that $L_0$ behavior is the same for all roles: here, this means that the probability that $L_0$ Hiders and $L_0$ Seekers select the most frequent strategy-label is equal to $q$. Further, a level-$k$ model would entail that Seekers select the label that is the most frequent choice of co-players at the level below, whereas Hiders select the label that is the least frequent choice of co-players at the level below. More specifically, this means that the most frequent strategy-label will be selected by $L_1$ Hiders and $L_1$ Seekers with probability zero and one, respectively. Similarly, iterated best responses are implemented by higher-level types and, hence, here both $L_2$ Hiders and $L_2$ Seekers will select the most frequent strategy-label with probability zero.

In brief, while particular hypotheses about the distribution of choices across conditions reflect the extent of subjects’ strategic sophistication (i.e., the distribution of types), for the purposes of this study we limit ourselves to verifying whether the effect of the frequency of exposure from Study 1 was nonstrategic: if it were, then we should observe the same choice distributions across the four conditions. By contrast, if participants use such frequency of exposure in a boundedly rational manner, then they will be more likely to select prominent strategy-labels in the Coordinate (or Pick) condition than in the Seek condition; moreover, they will be more likely to select prominent strategy-labels in the Seek condition than in the Hide condition.

**Results.** Table 1 presents (per-subject) mean choices in each of the four conditions, given a classification of the strategy-labels based on the NGRAM metric. For the distribution of individual-level choices in each of the 10 games, please refer to Figure 2A in the Appendix.

---

16 In fact, $L_1$ Hiders who expect a plurality (i.e., a fraction $q$) of their opponents to select the most prominent label will want to avoid that option. On the other hand, $L_1$ Seekers who expect a plurality of their opponents to select the most prominent label will want to choose that option.
Table 1 – (Per-subject) mean choice, given a classification of the strategy-labels based on the NGram metric; in brackets is the standard deviation.¹⁷

<table>
<thead>
<tr>
<th>Choice by frequency of exposure</th>
<th>Coordinate</th>
<th>Pick</th>
<th>Seek</th>
<th>Hide</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strategy-label with highest NGram metric is chosen ([f_H]), %</td>
<td>46.19 (.1464)</td>
<td>48.46 (.1646)</td>
<td>43.64 (.1556)</td>
<td>36.86 (.1548)</td>
</tr>
<tr>
<td>Strategy-label with middling NGram metric is chosen ([f_M]), %</td>
<td>48.10 (.1596)</td>
<td>38.21 (.1211)</td>
<td>34.09 (.1661)</td>
<td>32.85 (.1202)</td>
</tr>
<tr>
<td>Strategy-label with lowest NGram metric is chosen ([f_L]), %</td>
<td>5.71 (.0914)</td>
<td>13.33 (.1675)</td>
<td>22.27 (.2044)</td>
<td>30.29 (.2121)</td>
</tr>
<tr>
<td>Total, %</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Total # triplets (1,600) [\text{i.e., no. of triplets} \times \text{subjects}]</td>
<td>420</td>
<td>390</td>
<td>440</td>
<td>350</td>
</tr>
<tr>
<td>Subjects (160)</td>
<td>42</td>
<td>39</td>
<td>44</td>
<td>35</td>
</tr>
</tbody>
</table>

By giving a quick glance at the table the reader will easily notice that the distribution of choices varies with each condition. In particular, consistently with our prediction, the most prominent option (i.e., the strategy-label with the highest NGram metric) is selected less and less often when moving from Pick to Seek to Hide conditions: a Kruskal-Wallis test on the choice of such prominent strategy-labels confirms statistically significant differences across the four conditions (in order to satisfy the assumption of independence of observations, this test – like the following ones – is conducted on the sample of per-subject mean choices; \(N = 160\) obs., \(\chi^2_3 = 10.477, p = 0.014, \text{two-tailed}\)). On the other hand, it is noteworthy that the least prominent option (i.e., the strategy-label with the lowest NGram metric) is selected more and more often when moving from Coordinate to Pick to Seek to Hide conditions: a Kruskal-Wallis test on the choice of such “non-prominent” strategy-labels confirms statistically significant differences across the four conditions (\(N = 160\) obs., \(\chi^2_3 = 34.353, p = 0.000, \text{two-tailed}\)).

¹⁷ Given that the number of subjects in Hide was less than in Seek – for the mere purpose of calculating Seekers’ payoffs – nine participants in the former condition were matched with two Seekers (but received compensation for playing with either one, at random).
Further, following on from the above discussion of level-\( k \) reasoning, we will treat participants in the \textit{Pick} condition as control subjects – by virtue of being fully nonstrategic – and proceed to test our directional hypotheses via simple pairwise comparisons (using the sample of per-subject mean choices). We begin by contrasting choice behavior in the Pick and Coordinate conditions: a one-tailed test allows us to examine the alternative hypothesis that the prominent option (i.e., the strategy-label with the highest NGRAM metric) is selected strictly less often in Coordinate than in Pick.\(^{18}\) In short, a one-tailed Wilcoxon-Mann-Whitney test does not reject the null hypothesis of no difference: more explicitly, there is no evidence of a significant decrease in the choice of the most prominent option when moving from \textit{Pick} to \textit{Coordinate} (\( N = 81 \) obs., \( Z = -1.008, p = 0.1567 \)). This is consistent with our prediction.

Next, we find that the most prominent option is selected marginally less in \textit{Seek} than in \textit{Pick} (\( N = 83 \) obs., \( Z = 1.440, p = 0.074 \), under a one-tailed Wilcoxon-Mann-Whitney test);\(^{19}\) also, we find that the most prominent option is selected significantly less in \textit{Hide} than in \textit{Pick} (\( N = 74 \) obs., \( Z = 3.101, p = 0.000 \), under a one-tailed Wilcoxon-Mann-Whitney test). Besides, the most prominent option is selected significantly less in \textit{Hide} than in \textit{Seek} (\( N = 79 \) obs., \( Z = 1.978, p = 0.023 \), under a one-tailed Wilcoxon-Mann-Whitney test):\(^{20}\) this appears to suggest that Hiders tend to avoid strategies associated with highly popular linguistic labels.

In summary, consistently with our predictions, the above tests indicate that there are no differences between \textit{Coordinate} and \textit{Pick} conditions, whereas prominent options are selected less and less often when moving from \textit{Pick} to \textit{Seek} to \textit{Hide} conditions. Yet, we note that since the above tests are conducted on the sample of (per-subject) mean choices, they do not account for differences in individual-level responses across games, nor do they account for any effect of option ordering; moreover, the above tests do not account for differences in the relative magnitude of the NGRAM metrics across games. For these reasons, we corroborated our

\(^{18}\) Recall that level-\( k \) reasoning implies that subjects will be \textit{weakly more} likely to select the most frequent strategy-label in the Coordinate condition than in the Pick condition. Because of the weak inequality, here we test the null hypothesis of no difference against the alternative hypothesis that the prominent option is selected \textit{strictly less} often in Coordinate than in Pick.

\(^{19}\) Note that here we test the null hypothesis of no difference against the alternative hypothesis that subjects are \textit{strictly more} likely to select prominent strategy-labels in the Pick condition than in the Seek condition (as directly implied by level-\( k \) reasoning).

\(^{20}\) Here we test the null hypothesis of no difference against the alternative hypothesis that subjects are \textit{strictly more} likely to select prominent strategy-labels in the Seek condition than in the Hide condition (as directly implied by level-\( k \) reasoning).
analysis by running alternative-specific conditional logistic (asclogit) regressions that account for the relevant NGRAM metric of each option – in each game – as well as for ordering effects. In what follows we summarize the main findings while we refer the reader to the Appendix for the full regression tables.

In short, our asclogit regressions confirm a positive and significant effect of exposure, as measured by our NGRAM metrics, across treatments: this means that the higher the frequency of exposure to a label, the more likely it is for the associated strategy to be selected, regardless of its position. In particular, when contrasting choice behavior in *Pick* and *Coordinate*, our asclogit regression reveals no significant difference in the relative impact of the NGRAM metric across the two conditions (see regressor *FOC* in the left panel of Table 2A in the Appendix): in other words, we find no difference in the probability of choosing prominent options between the *Pick* and *Coordinate* conditions. On the other hand, when contrasting choice behavior in *Pick* and *Seek*, we find a mildly significant difference in the relative impact of the NGRAM metric on choice across the two conditions: that is, we find that an increase in the NGRAM metric is more likely to drive the choices of subjects in *Pick* than in *Seek* (coef. = -0.1891, \( z = -1.85 \), \( p = 0.065 \), under a two-tailed asclogit regression conducted on the full sample of individual observations, with standard errors adjusted for clustering on 83 subjects; see regressor *FOC* in the right panel of Table 2A). Moreover, when contrasting choice behavior in *Seek* and *Hide*, we find a significant difference in the relative impact of the NGRAM metric on choice across the two conditions: that is, we find that an increase in the NGRAM metric is more likely to drive the choices of subjects in *Seek* than in *Hide* (coef. = -0.2145, \( z = -2.14 \), \( p = 0.033 \), under a two-tailed asclogit regression conducted on the full sample of individual observations, with standard errors adjusted for clustering on 79 subjects; see regressor *FOC* in the right panel of Table 3A). Please refer to the Appendix for an extended commentary.

To conclude, the data fully support our hypotheses: prominent options are selected less and less often when moving from *Pick* to *Seek* to *Hide* conditions; such data patterns suggest that subjects behave in a boundedly rational manner consistent with level-*k* reasoning. We can now proceed to the discussion of our next study, which was designed to further explore the strategic use of strategy-labels in coordination games. To that end, in what follows we shall present a few novel conditions, in which we varied the cultural/linguistic makeup of the subject pool as well as participants’ knowledge of the counterpart’s country of residence (and thus cultural background).
Indeed, this type of shared cultural knowledge facilitated by frequency of exposure is an implicit feature of Shelling’s informal definition of prominence.

Study 3

Demographics. The subject pool for Study 2 consisted of 80 individuals, of which half were US residents and half were UK residents. All participants were recruited from the Prolific Academic website, simultaneously. In the US-residents sample, the average participant was 29 years old and 75% of subjects were male. In the UK-residents sample, the average participant was 31 years old and 63% of subjects were male. As with our previous studies, all subjects received a 0.5 GBP participation fee, besides any additional payoffs resulting from their choices (as described below).21

Design. Study 3 involved the same 2-player pure coordination structure as in Study 1 (see the matching game of Figure 1 above), except that we used a different set of options, as reported in Figure 4 below. Furthermore, in Study 3 we manipulated the information provided to subjects in regards to their partners’ country of residence. Specifically, each subject in both the US and UK samples was assigned to one of the following three information conditions.

a. “No-k”: Participants in this condition received no information about their partners’ residence (hence, this condition is identical to Study 1, except that we used different strategy-labels). The instructions in this condition read: “Please choose one option. Each of you and your partner receive £0.10 if you both choose the same option, £0 otherwise.”

b. “k-UK”: Participants in this condition were told that their partner resided in the UK. Specifically, they were shown the following message: “Please choose one option. Each of you and your partner receive £0.10 if you both choose the same option, £0 otherwise. Your partner is a Prolific worker who resides in the UK. Your partner may or may not know where you reside.”

c. “k-US”: Participants in this condition were told that their partner resided in the US. Specifically, they were shown the following message: “Please choose one option. Each

21 Average earnings were GBP 0.46 and 0.38 for the US and UK samples, respectively (not counting subjects’ participation fees).
of you and your partner receive £0.10 if you both choose the same option, £0 otherwise.
Your partner is a Prolific worker who resides in the US. Your partner may or may not know where you reside."

In summary – in both the US and UK samples – each subject was assigned to one of the above three conditions, implying that our study features a total of six sub-conditions: 1) US participants without additional knowledge about their partner; 2) UK participants without additional knowledge about their partner; 3) US participants knowing that they are matched with a US participant; 4) US participants knowing that they are matched with a UK participant; 5) UK participants knowing that they are matched with a US participant; 6) UK participants knowing that they are matched with a UK participant.

The order of the 10 games was randomized across subjects. By contrast, the order of the three options (in a given game) was determined randomly prior to the experiment, but was identical across subjects; the order of the options is reported in Figure 4 below (i.e., option X, option Y, option Z). Note that we generated this list of triplets so that the option with the highest frequency of exposure would vary between the American and British English vocabularies (as measured by the NGRAM metrics for the American and British English Google Books corpora, respectively). For example, in the first game “curry” has the highest NGRAM metric in British English but the lowest one in American English; conversely, “chili” has the highest NGRAM metric in American English and the lowest one in British English.

The psychological mechanism we outlined earlier on here suggests that, if one is informed that the assigned partner resides in the same country (and thus shares the same cultural/linguistic background), then one will more likely choose the strategy associated with the most “prominent label”, as measured by the NGRAM metric pertaining to one’s own vocabulary. More precisely – compared to the condition in which one is informed that the assigned partner resides in the same country – we predict that the prominent label is selected weakly less often if one is unaware of the partner’s country of residence, and strictly less often if one is informed that the partner resides in a different country.
Figure 4 – The option sets for Study 3. The blue, maroon, and green cells in the first (second) row below each strategy-label indicate the option with the relatively highest, middling and lowest US (UK) NGRAM metric, respectively.
**Results.** Table 2 presents (per-subject) mean choices in each of the six sub-conditions, given a classification of the strategy-labels based on the relevant NGRAM metric. For the distribution of individual-level choices in each of the 10 games, please refer to Figure 3A (**US sample**) and Figure 4A (**UK sample**) in the Appendix.

<table>
<thead>
<tr>
<th>Choice by frequency of exposure</th>
<th>US sample</th>
<th>UK sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No-(k)</td>
<td>(k)-(UK)</td>
</tr>
<tr>
<td>Strategy-label with <em>highest</em> NGRAM metric is chosen ([f_{H}]), %</td>
<td>47.89</td>
<td>(.1812)</td>
</tr>
<tr>
<td>Strategy-label with <em>middling</em> NGRAM metric is chosen ([f_{M}]), %</td>
<td>31.05</td>
<td>(.1663)</td>
</tr>
<tr>
<td>Strategy-label with <em>lowest</em> NGRAM metric is chosen ([f_{L}]), %</td>
<td>21.06</td>
<td>(.1559)</td>
</tr>
<tr>
<td>Total, %</td>
<td>100</td>
<td>(.1828)</td>
</tr>
</tbody>
</table>

| Total # triplets | i.e., no. of triplets * subjects | 190 | 100 | 110 | 180 | 120 | 100 |

| Subjects | 19 | 10 | 11 | 18 | 12 | 10 |

**Table 2** – (Per-subject) mean choice, given a classification of the strategy-labels based on the relevant NGRAM metric (American English corpus for **US sample** and British English corpus for **UK sample**); in brackets is the standard deviation.

By giving a glance at Table 2 above the reader will quickly notice that the distribution of choices varies with each sub-condition.\(^{22}\) We proceed to test our directional hypotheses via simple pairwise comparisons. We begin by analyzing the **US sample**. First, we find that the most prominent option (i.e., the strategy-label with the highest **US** NGRAM metric) is selected marginally more in \(k\)-**US** than in No-\(k\) (in order to satisfy the assumption of independence of

\(^{22}\) A Kruskal-Wallis test on the choice of the prominent option (i.e., the strategy-label with the highest NGRAM metric pertaining to the relevant vocabulary) indicates mildly significant differences across the three sub-conditions of the **US sample** (\(N = 40\) obs., \(\chi^2 = 5.931, p = 0.051\), two-tailed). Similarly, a Kruskal-Wallis test indicates mildly significant differences across the three sub-conditions of the **UK sample** (\(N = 40\) obs., \(\chi^2 = 4.837, p = 0.089\), two-tailed).
observations, this test – like the other non-parametric tests – is conducted on the sample of per-subject mean choices; \(N = 30\) obs., \(Z = -1.382, p = 0.083\), under a one-tailed Wilcoxon-Mann-Whitney test). Second, as we predicted, we find that the most prominent option is selected more often in \(k-US\) than in \(k-UK\) \((N = 21\) obs., \(Z = -2.250, p = 0.012\), under a one-tailed Wilcoxon-Mann-Whitney test). Besides, we find that the most prominent option is selected more often in \(No-k\) than in \(k-UK\) \((N = 29\) obs., \(Z = 1.657, p = 0.048\), under a one-tailed Wilcoxon-Mann-Whitney test).

We now turn to analyze the \(UK\) sample. First, we find that there is no significant difference in the choice of the prominent option (i.e., the strategy-label with the highest \(UK\) NGRAM metric) between \(k-UK\) and \(No-k\) \((N = 30\) obs., \(Z = -1.202, p = 0.114\), under a one-tailed Wilcoxon-Mann-Whitney test). Second, as we predicted, we find that the most prominent option is selected more often in \(k-UK\) than in \(k-US\) \((N = 22\) obs., \(Z = 2.104, p = 0.017\), under a one-tailed Wilcoxon-Mann-Whitney test). Besides, we find that the most prominent option is selected marginally more in \(No-k\) than in \(k-US\) \((N = 28\) obs., \(Z = 1.416, p = 0.078\), under a one-tailed Wilcoxon-Mann-Whitney test).

Like before, we note that these non-parametric tests provide a convenient overview of the main patterns in our aggregate data, however they do not control for a number of relevant variables (e.g., differences in individual-level responses across games, option ordering, and relative magnitude of the NGRAM metrics across games). Thus, as with our previous studies, we corroborated our findings by running some alternative-specific conditional logistic regressions (all regression tables in the Appendix). In short, such regressions provide further evidence in support of the hypothesis that, if one is informed that the assigned partner resides in the same country, one will more likely choose the strategy associated with the most “prominent label” (as measured by the NGRAM metric pertaining to one’s own vocabulary).

In particular – in the case of the \(US\) sample – when contrasting choice behavior in \(k-US\) and \(k-UK\), we find a strongly significant difference in the relative impact of the relevant NGRAM metric on choice: that is, we find that an increase in the NGRAM (\(US\)) metric is significantly more likely to drive the choices of subjects in \(k-US\) than in \(k-UK\) \((\text{coef.} = -4.316, z = -2.93, p = 0.003\), under a two-tailed asclogit regression conducted on the full sample of individual observations, with standard errors adjusted for clustering on 21 subjects; see regressor \(FOC\) in the left panel of Table 4A). This means that US residents are more likely to choose
strategies relating to highly popular labels (in the *American* English vocabulary) when they are informed that their assigned partner resides in the US rather than in the UK.

Similarly – in the case of the *UK sample* – when contrasting choice behavior in *k-UK* and *k-US*, like before we find a strongly significant difference in the relative impact of the relevant NGRAM metric on choice: that is, we find that an increase in the NGRAM (*UK*) metric is significantly more likely to drive the choices of subjects in *k-UK* than in *k-US* (coef. = -4.218, *z* = -2.94, *p* = 0.003, under a two-tailed asclogit regression conducted on the full sample of individual observations, with standard errors adjusted for clustering on 22 subjects; see regressor *FOC* in the right panel of Table 4A). Again, this means that UK residents are more likely to choose strategies relating to highly popular labels (in the *British* English vocabulary) when they are informed that their assigned partner resides in the UK rather than in the US.

Finally, we conclude this section by testing more directly for the relationship between one’s strategic use of prominence (i.e., frequency of exposure) and one’s knowledge that the assigned partner is alike. To do so, we coded a “knowledge” ordinal variable as follows: knowledge takes on value 1 if the subject is informed that the assigned partner resides in a different country; it takes on value 2 if the subject receives no information about the assigned partner’s country of residence; it takes on value 3 if the subject is informed that the assigned partner resides in the same country. We then coded a “prediction” variable, which takes on value 1 if a subject selects the most prominent label (as measured by the NGRAM metric pertaining to one’s own vocabulary) and takes on value 0 otherwise. Given that, we regressed the prediction variable on the knowledge level: the results are striking. In the *US sample* we find a positive and strongly significant effect of knowledge on prediction (coef. = .4529, *z* = 2.65, *p* = 0.008, under a two-tailed standard logistic regression with standard errors clustered on 40 subjects). Unsurprisingly, we find a similar effect in the *UK sample* (coef. = .3851, *z* = 2.44, *p* = 0.015, under a two-tailed standard logistic regression with standard errors clustered on 40 subjects). This confirms that subjects do use prominent labels strategically: the more they know that their assigned partner is alike, the more likely they are to select the prominent strategy-label.
Conclusion

We have investigated games with natural language labels, and measured the prominence of strategies by how often their labels occur in natural language (specified by the Google Books NGRAM frequencies). In the first study we found that the frequency with which labels are mentioned in natural language accurately predict the likelihood that participants select the strategies corresponding to the labels, in matching games (i.e., pure coordination games). Our second study was designed to verify whether subjects use “prominent labels” (in the sense of the above definition) strategically rather than naively, such as an automatic response. To do so, we contrasted the behavior of subjects participating in a matching game (“Coordinators”) with the behavior of subjects facing three alternative scenarios (which each scenario featuring exactly the same list of labels); namely, the case of “Pickers”, “Hiders”, and “Seekers”. Our data show that Pickers were about as likely to select the most prominent label as were Coordinators; by contrast, Hiders were less likely to select the most prominent label than were Seekers and, in turn, Seekers were less likely than Pickers. This pattern suggests a boundedly rational use of strategy-labels. Finally, we designed our third study to corroborate our results on the strategic use of strategy-labels in matching games: to that end, we varied the participants’ knowledge of the counterpart’s country of residence. Our data show that subjects were less likely to rely on the frequency of exposure as a means to guiding their behavior when they knew that their counterpart resided in a different country and thus had a different cultural/linguistic background.

This paper has proposed a psychologically grounded characterization of the notion of prominence as a criterion for equilibrium selection in strategic games. In brief, frequency of exposure is one of the fundamental determinants of the mental accessibility of concepts, and previous research has shown that individuals often rely on such frequency of exposure to guide their behavior in a variety of non-strategic tasks (e.g., memory, language, and decision tasks): in fact, identifying the most popular linguistic label is automatic and therefore effortless. This suggests that – in the context of strategic games – subjects may perceive as particularly prominent those strategies that are denoted by highly popular (linguistic) labels; hence, upon facing a strategic problem, subjects may be naturally drawn to such prominent strategy-labels when considering alternative courses of action. Now, if players “project” their mental categories onto others, they will realize that others too can identify the same, prominent strategy-labels (and especially so when they know that their partners come from the same linguistic environment)!
Since these mental processes imply almost no cognitive burden, then relying on frequency of exposure to guide one’s decisions turns out to be an efficient decision rule (in terms of time and effort) for coordination and discoordination problems.

To conclude, our set of studies has illustrated how strategic behavior can be accurately predicted in games with natural language labels. For the purposes of this paper we have limited our focus to coordination and discoordination games, but our approach may be extended to other classes of games (so long as strategies are denoted by linguistic labels). Another direction for further research involves the application of natural language statistics and corpus analysis to analyze the semantic similarity between strategy-labels, in such a way as to account for cases in which a non-popular linguistic label may stand out as salient by virtue of being dissimilar from the rest of the options.\textsuperscript{23}

\textsuperscript{23} This case does not occur in our experiments because each member of a triplet belongs to the same category (i.e., either countries or food), and hence our strategy-labels are fairly homogenous. That said, we note that previous research has shown that dissimilarity can often influence choice behavior in matching games where strategies are identified by diverse objects or shapes (see, among the others, Bacharach et al., 1997).
Appendices

Appendix - Study 1

Figure 1A below shows the frequency distributions of individual-level choices in each of the (20) games. There – for ease of exposition – we have rearranged the order of the options in each game, in such a way that the most prominent option (i.e., the one with the highest NGRAM metric, and thus our predicted choice) is always shown at the bottom of the bar.

**Figure 1A** – Frequency distributions of individual-level choices. The left and right panels refer to Versions A and B of Study 1, respectively. The blue, maroon, and green bars indicate the frequency with which the option with the relatively highest, middling and lowest NGRAM metric, respectively, was chosen.
In what follows we report the output of an alternative-specific conditional logistic ("asclogit") regression.\textsuperscript{24} There, the likelihood of a subject’s choice may depend on: i) “case-specific regressors”, which reflect characteristics that are common to all alternatives in a group (e.g., subjects assigned to the same version of Study 1 belong to the same group); ii) “alternative-specific regressors”, which reflect particular characteristics of the alternatives one faces in a certain problem (e.g., the frequency of exposure to a label);\textsuperscript{25} iii) characteristics that are specific to the interaction of i and ii.

The model in Table 1A is characterized by a central, upper section (presenting alternative-specific regressors) and a bifurcated, lower section (presenting case-specific regressors). Let’s start from the latter. Once again, case-specific regressors refer to characteristics that are common to all alternatives in a group, such as the version of the study one is assigned to: note that case-specific coefficients are interpreted as parameters of an ordinary multinomial logit model, against the base category (which in Table 1A is specified directly above the bifurcation).

On the other hand, alternative-specific regressors refer to attributes that are specific to an option, as opposed to being specific to the whole triplet of options. Alternative-specific coefficients measure the impact of a change in an attribute (on the likelihood of choosing an option featuring that attribute). So, if a regressor has a positive coefficient, then an increase in that attribute (e.g., an increase in the relative frequency of exposure to a label) implies that the corresponding strategy is selected more often than the other two strategies in the triplet, whichever they are.\textsuperscript{26}

More specifically, the model in Table 1A presents: a continuous variable ("FO") measuring the relative frequency of exposure to a strategy-label; a dummy ("V") indicating whether an option is shown to a participant in Version B; an interaction variable ("FOV").\textsuperscript{27}

\textsuperscript{24} Note that conditional logistic regressions differ from regular logistic analysis as the data are grouped (in such a way that all the observations in a group have some common characteristic) and, therefore, the likelihood is calculated relative to each group.

\textsuperscript{25} Note that unlike multinomial logit models, where all the independent variables are case-specific, the asclogit model allows for both case- and alternative-specific predictors. For a general account of the asclogit model, see Cameron and Trivedi (2010, pp. 503-511).

\textsuperscript{26} It is now clear why the output of the asclogit model consists of a unique coefficient for each alternative-specific regressor, without an explicit base category.

\textsuperscript{27} Note that regressor FOV is alternative specific in that it captures how the impact of exposure varies with each version of the study.
<table>
<thead>
<tr>
<th>Choice of option</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Relative frequency of exposure to a label (&quot;FO&quot;)</td>
<td>.891***</td>
</tr>
<tr>
<td></td>
<td>(.142)</td>
</tr>
<tr>
<td>Relative frequency of exposure * Version (&quot;FOV&quot;)</td>
<td>-.359***</td>
</tr>
<tr>
<td></td>
<td>(.090)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Center vs. Top</th>
<th>Bottom vs. Top</th>
</tr>
</thead>
<tbody>
<tr>
<td>Version B (&quot;V&quot;)</td>
<td></td>
</tr>
<tr>
<td>.789***</td>
<td>.303</td>
</tr>
<tr>
<td>(.252)</td>
<td>(.233)</td>
</tr>
<tr>
<td>Constant</td>
<td></td>
</tr>
<tr>
<td>-.1772***</td>
<td>-1.084***</td>
</tr>
<tr>
<td>(.430)</td>
<td>(.375)</td>
</tr>
</tbody>
</table>

Log pseudolikelihood | -899.663
Obs.                 | 910

Table 1A – Study 1: alternative-specific conditional logistic regression coefficients, with clustered standard errors in brackets (*, **, and *** indicate $p<0.10$, $p<0.05$ and $p<0.01$, respectively, for the relevant Z Statistic, two-tailed tests).

We note that FO is positive and strongly significant: this means that if an option occurs more frequently in natural language, then it is more likely chosen – across versions – regardless of location (i.e., top, center, or bottom of the list).\(^{28}\) Also, note that FOV is negative, indicating that the impact of exposure in Version B is relatively smaller than in Version A. Two more comments are in order. Firstly, the constant terms reported in the left and right branches respectively indicate the general desirability of the middle and bottom options (against the top option) due to unmeasured attributes of the alternatives: note that they are both negative, which reflects the greater desirability of the top options. Secondly, the main-effects coefficient for the version regressor reveals that participants in Version B are significantly more likely to pick the middle option against the top option (as indicated by the leftmost coefficient for V).

\(^{28}\) We stress that this contradicts the prescriptions of traditional game theory (Harsanyi and Selten, 1988), according to which – in matching games – one should always play the fully mixed equilibrium.
Appendix - Study 2

Figure 2A below shows the frequency distributions of individual-level choices in each of the 1-10 games and for each of the 1-4 conditions. Like before – for ease of exposition – we have rearranged the order of the options, in such a way that the most prominent option (i.e., the one with the highest NGRAM metric) is always shown at the bottom of the bar.

Figure 2A – Frequency distributions of individual-level choices, in each of the 1-10 games and for each of the 1-4 conditions. NOTE: Conditions 1-4 refer to “Coordinate”, “Pick”, “Seek”, and “Hide”, respectively. The blue, maroon, and green bars indicate the frequency with which the option with the relatively highest, middling and lowest NGRAM metric, respectively, was chosen.

By giving a glance at the graph one can easily notice that the distribution of choices varies with each condition. Moreover, there is a trend whereby the choice frequency of the option with the highest NGRAM metric typically decreases when moving from Pick to Hide conditions (i.e., from bar 2 to bar 4).
We now present a few alternative-specific conditional logistic (asclogit) regressions. In particular, the left panel in Table 2A uses the full sample of option triplets (i.e., ten per subject) from the *Pick* and *Coordinate* conditions, whereas the right panel in Table 2A uses the full sample of triplets from the *Pick* and *Seek* conditions. Furthermore, the left panel in Table 3A uses the full sample of triplets from the *Pick* and *Hide* conditions, whereas the right panel in Table 3A uses the full sample of triplets from the *Seek* and *Hide* conditions. All the regressions involve the following predictors: a continuous variable (“FO”) measuring the relative frequency of exposure to a strategy-label; a dummy (“C”) indicating whether an option is shown to a participant in the alternative condition (as specified in the table notes); an interaction variable (“FOC”).

<table>
<thead>
<tr>
<th>Choice of option</th>
<th>Pick &amp; Coordinate</th>
<th>Pick &amp; Seek</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relative frequency of exposure to a label (&quot;FO&quot;)</td>
<td>.637*** (.161)</td>
<td>.915*** (.267)</td>
</tr>
<tr>
<td>Relative frequency of exposure * Condition (&quot;FOC&quot;)</td>
<td>-.050 (.104)</td>
<td>-.189* (.102)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Condition (&quot;C&quot;)</th>
<th>Center vs. Top</th>
<th>Bottom vs. Top</th>
<th>Center vs. Top</th>
<th>Bottom vs. Top</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-.948*** (.212)</td>
<td>-.853*** (.212)</td>
<td>.209 (.183)</td>
<td>.043 (.159)</td>
</tr>
<tr>
<td>Constant</td>
<td>-.357*** (.115)</td>
<td>-.098 (.110)</td>
<td>-.357*** (.115)</td>
<td>-.098 (.110)</td>
</tr>
</tbody>
</table>

Log pseudolikelihood: -746.093 (-850.131)  
Obs. 810 (830)

**Table 2A** – Study 2: alternative-specific conditional logistic regression coefficients, with clustered standard errors in brackets (*, **, and *** indicate \( p<0.10 \), \( p<0.05 \) and \( p<0.01 \), respectively, for the relevant Z Statistic, two-tailed tests). LEFT panel: regressor *C* takes on value 0 if a subject is assigned to *Pick*; it takes on value 1 if a subject is assigned to *Coordinate*. RIGHT panel: regressor *C* takes on value 0 if a subject is assigned to *Pick*; it takes on value 1 if a subject is assigned to *Seek*. 
<table>
<thead>
<tr>
<th>Choice of option</th>
<th>Pick &amp; Hide</th>
<th>Seek &amp; Hide</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relative frequency of exposure to a label (&quot;FO&quot;)</td>
<td>.940*** (.171)</td>
<td>.991*** (.350)</td>
</tr>
<tr>
<td>Relative frequency of exposure * Condition (&quot;FOC&quot;)</td>
<td>-.201*** (.053)</td>
<td>-.214** (.100)</td>
</tr>
<tr>
<td>Condition (&quot;C&quot;)</td>
<td>Center vs. Top .535*** (.174) Bottom vs. Top .169 (.168)</td>
<td>Center vs. Top .325* (.192) Bottom vs. Top .126 (.172)</td>
</tr>
<tr>
<td>Constant</td>
<td>Center vs. Top -.357*** (.115) Bottom vs. Top -.098 (.110)</td>
<td>Center vs. Top -.148 (.141) Bottom vs. Top -.054 (.115)</td>
</tr>
<tr>
<td>Log pseudolikelihood</td>
<td>-767.819</td>
<td>-846.129</td>
</tr>
<tr>
<td>Obs.</td>
<td>740</td>
<td>790</td>
</tr>
</tbody>
</table>

Table 3A – Study 2: alternative-specific conditional logistic regression coefficients, with clustered standard errors in brackets (*, **, and *** indicate $p<0.10$, $p<0.05$ and $p<0.01$, respectively, for the relevant Z Statistic, two-tailed tests). LEFT panel: regressor $C$ takes on value 0 if a subject is assigned to Pick; it takes on value 1 if a subject is assigned to Hide. RIGHT panel: regressor $C$ takes on value 0 if a subject is assigned to Seek; it takes on value 1 if a subject is assigned to Hide.

Note that the strong significance of $FO$ in all regressions confirms a positive effect of exposure (as measured by our NGRAM metrics) across treatments: in other words, the higher the frequency of exposure to a label, the more likely it is for the associated strategy to be selected, regardless of its position. Furthermore, note that $FOC$ is non-significant in the left panel of Table 2A: this means that there is no significant difference in the relative impact of the NGRAM metric between Pick and Coordinate conditions. Moreover, note that $FOC$ is significant in all the other regressions: this means that prominent options are selected less often when moving from Pick to Seek, from Pick to Hide, and from Seek to Hide conditions.
Appendix - Study 3

Figures 3A and 4A below show the frequency distributions of individual-level choices (in each of the 1-10 games and for each of the 1-3 conditions) in the US and UK samples, respectively. Note that – for ease of exposition – we have rearranged the order of the options for each game and each sample, in such a way that the option with the highest NGRAM metric (according to the relevant vocabulary) is always shown at the bottom of the bar.

Figure 3A – Frequency distributions of individual-level choices – for each of the 1-10 games and for each of the 1-3 information conditions – US sample only. NOTE: Conditions 1-3 refer to “No-k”, “k-UK”, and “k-US”, respectively. The blue, maroon, and green bars indicate the frequency with which the option with the relatively highest, middling and lowest NGRAM (US) metric, respectively, was chosen.
Figure 4A – Frequency distributions of individual-level choices – for each of the 1-10 games and for each of the 1-3 information conditions – UK sample only. NOTE: Conditions 1-3 refer to “No-k”, “k-UK”, and “k-US”, respectively. The blue, maroon, and green bars indicate the frequency with which the option with the relatively highest, middling and lowest NGRAM (UK) metric, respectively, was chosen.

A close look to the graphs reveals that the distribution of choices varies with the information condition. Moreover, two similar trends appear to emerge across the US and UK samples. Specifically, in the US sample (Figure 3A) the most prominent option for US participants (the one with the highest NGRAM value in the American corpus) is typically the most likely to be selected. On the other hand, in the UK sample (Figure 4A) the most prominent option for UK participants (the one with the highest NGRAM value in the British corpus) is often more likely selected.

We now present some alternative-specific conditional logistic (asclogit) regressions. The left panel in Table 4A refers to the US sample and uses the full set of option triplets (i.e., ten per subject) from the k-UK and k-US conditions. On the other hand, the right panel in Table 4A
refers to the *UK sample* and uses the full set of option triplets from the same information conditions (i.e., *k-US* and *k-UK*). Both regressions involve the following predictors: a continuous variable (“FO”) measuring the relative frequency of exposure to a strategy-label (i.e., the NGRAM metric relative to the *American* corpus in the left panel, and the NGRAM metric relative to the *British* corpus in the right panel); a dummy (“C”) indicating whether an option is shown to a participant in the alternative condition (as specified in the table notes); an interaction variable (“FOC”).

<table>
<thead>
<tr>
<th>Choice of option</th>
<th><em>k-US</em> &amp; <em>k-UK</em></th>
<th><em>k-US</em> &amp; <em>k-UK</em></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>(US sample)</em></td>
<td><em>(UK sample)</em></td>
</tr>
<tr>
<td>Relative frequency of exposure to a label (“FO”)</td>
<td>4.479*** (.747)</td>
<td>1.841** (.861)</td>
</tr>
<tr>
<td>Relative frequency of exposure * Condition (“FOC”)</td>
<td>-4.316*** (1.471)</td>
<td>-4.218*** (1.436)</td>
</tr>
<tr>
<td>Condition (“C”)</td>
<td>Center vs. Top</td>
<td>Bottom vs. Top</td>
</tr>
<tr>
<td></td>
<td>.193 (.550)</td>
<td>.478 (.508)</td>
</tr>
<tr>
<td>Constant</td>
<td>-.491 (.395)</td>
<td>-1.078*** (.316)</td>
</tr>
<tr>
<td>Log pseudolikelihood</td>
<td>-203.014</td>
<td>-220.462</td>
</tr>
<tr>
<td>Obs.</td>
<td>210</td>
<td>220</td>
</tr>
</tbody>
</table>

*Table 4A* – Study 3: alternative-specific conditional logistic regression coefficients, with clustered standard errors in brackets (*, **, and *** indicate *p*<0.10, *p*<0.05 and *p*<0.01, respectively, for the relevant Z Statistic, two-tailed tests). LEFT panel: regressor *C* takes on value 0 if a subject is assigned to *k-US*; it takes on value 1 if a subject is assigned to *k-UK*. RIGHT panel: regressor *C* takes on value 0 if a subject is assigned to *k-UK*; it takes on value 1 if a subject is assigned to *k-US*.

Note that *FO* is positive and significant in both regressions: this means that the more prominent is an option (i.e., the strategy-label with the highest NGRAM metric pertaining to the relevant
vocabulary), the more likely it is chosen. Further, note that FOC is negative and strongly significant in both regressions: this means that prominent options are selected less often when a subject is informed that her assigned partner resides in a different country than hers.

References


Cameron, A. Colin and Pravin K. Trivedi. 2010. Microeconometrics Using Stata, Revised Edition. College Station, TX: Stata Press.


